Temporal Dynamics of Cross-modal Affective Representations in Social and Object Categories

Tiziano Suran

Abstract

Affective features have been suggested to play a predominant role in social group representations. Although studies identified overlaps between affective processing regions and social groups representations, not much is known about the temporal dynamics of their interaction. In the present study, I used multivariate decoding of electroencephalography (EEG) data from an evaluative task to track affective representations as a function of stimulus semantic category. Pattern classifiers were trained at distinguishing between positive and negative valence concepts presented as words and then tested on the same concepts presented as pictures. In contrasting performance between social groups and objects, the results revealed a stronger affective decoding for social groups in both an early and late time windows, coinciding with delayed evaluative responses. The present findings provide initial evidence of the presence of categorical differences in affective temporal dynamics, and point to an increased complexity in social group affective/semantic representations.

1. Introduction

Human lesion studies suggest that social categories such as social groups might be represented in an independent brain network from non-social categories (Piretti et al., 2015; Rumiati, Carnaghi, Improta, Diez, & Silveri, 2014). Moreover, differently from nonsocial categories, the representation of social groups seems to give affective features a significantly greater weight. In a TMS study, we recently provided evidence that stimulating the inferior frontal gyrus, commonly associated with the processing of negative affective features, speeds up the categorization of negative social group names. That is, when participants were presented with negative social group names such as “butchers”, as opposed to the positive category of “musicians”, TMS speeded up categorization responses to the former, thus suggesting a link between affective processing and semantic processing (Suran, Rumiati, & Piretti, 2019). Moreover, using an affective priming paradigm, we found greater facilitation in evaluating social groups relative to objects following the presentation of semantically unrelated but affectively congruent primes. This suggests that the elaboration of affective information enhances the subsequent processing of social groups (Suran, Arcara, Piretti, & Rumiati, 2019). This set of findings is consistent with theories of semantic memory positing a central role of affective features in the representation of social groups (Lambon Ralph, Jefferies, Patterson, & Rogers, 2016; Mahon & Caramazza, 2011; Simmons & Barsalou, 2003). As for how this central role of affect may be implemented anatomically, previous neuroimaging studies identified activation in regions selectively responding only to affective information related to social groups, as well as areas responding to both social groups and nonsocial categories but showing an increased response to the former (Norris, Chen, Zhu, Small, & Cacioppo, 2004). Even within regions processing both types of concepts, it is possible that the greater weight of affective information for social groups may be also reflected in the temporal dynamics of its processing, an aspect that fMRI, given its relatively low temporal resolution may have been unable to capture. To overcome such limitation, in the present study we used the temporal resolution of electroencephalography (EEG) to study whether differences between social and non-social are present in the temporal dynamics of processing their affective features. Additionally, to control for possible confounds of modality-specific mechanisms, we investigated the neural correlates of affective processes common to both visual and lexical inputs.

According to one conceptualization of how semantic knowledge might be organized in the brain, affective features are independent from features that rely on other modalities (Lambon Ralph et al., 2016). A first characteristic of affective features is that they do not possess a dedicated input modality. This implies that, to study affect, one needs to rely on stimuli, such as words or pictures, which, despite eliciting the same affective processes, are also expected to rely on distinct modality-specific brain regions. For this reason, if one were to identify differences between two categories in affective features processing, it is often not possible to rule out the possibility that the difference is due to other properties of the stimulus, such as visual features in pictures that cannot be fully matched (Fairhall & Caramazza, 2013; Leonardelli, Fait, & Fairhall, 2019). By jointly using the information from two distinct input modalities, multivariate pattern analysis (MVPA) of neural information allows to overcome this limitation by individuating cognitive processes that are shared by different input modalities via cross-decoding (Grootswagers, Wardle, & Carlson, 2017; King & Dehaene, 2014). To this end, for each time point from the target word onset, a classifier was trained to distinguish valence-dependent neural patterns, and then tested on recognizing them in pictures representing the same concepts (King & Dehaene, 2014). Each classifier trained on words was tested on all time intervals from picture presentation, lasting up to 1 second post stimulus-onset. The resulting time generalization analysis allowed us to study the changes in valence-specific brain patterns, and to directly compare them between word and picture targets in the temporal domain (King & Dehaene, 2014). To this end, we asked participants to complete an evaluative task in which they were required to explicitly focus on the affective content of the stimuli while their brain responses were being recorded.

Affect-driven effects common to both words and pictures have been reported before in early and late time windows in studies adopting univariate analysis approaches. For example, singling-out negative information (associated with reactions to threat) can occur as early as in the 165-195 ms time window after target word onset (Zhang et al., 2014), while a later time window (~450 ms post-stimulus) has been argued to reflect the automatic processing of the polarity of affectively charged words (Zhang et al., 2014), the presence of an affective content of greater arousal (Hinojosa, Carretié, Valcárcel, Méndez-Bértolo, & Pozo, 2009), as well as the start of a more controlled, contextdependent processing of affective information (Cunningham, Espinet, Deyoung, & Zelazo, 2005). In studies using picture stimuli, the presence of valenced as opposed to affectively neutral content is found in an early (150-180 ms) time window, whereas the affective polarity of such stimuli, as for words, is evidenced by a later component (~450 ms, Zhu et al., 2015). If such temporal and activation similarities reflect the involvement of shared underlying processes, we expected to record modality-independent valence effects both in early (~150 ms) and later time windows (~450 ms). Based on the hypothesized interaction between the semantic category and affective processing, we expected an earlier and better modality-independent decoding when processing social compared to nonsocial categories at all valence processing stages.

2. Method

2.1 Participants

Twenty-one healthy participants (16 females, age range: 21 – 33 years) took part to the study for monetary compensation. The inclusion criteria consisted in speaking Italian as first language, self-reported right-handedness, and a minimum of 95% conformity in evaluating the target word stimuli according to the valence manipulation based on an online survey. The study protocol was approved and carried out in accordance with the recommendations of the local Ethics Committee, and in accordance with the Declaration of Helsinki. All subjects gave written informed consent prior to participating.

2.2 Materials

A total of 40 nouns of social groups (N = 20) and objects (N = 20) was selected from a larger database rated by a different sample of 12 subjects and used as target word stimuli. Each semantic category was equally split in a positive and negative subset of 10 elements each, differing significantly in their valence ratings (see Appendix 3). Within the positive and the negative valence stimuli, social group and object nouns were matched according to their average ratings of valence, arousal, familiarity, and length (see Table 1).

To exclude the possibility of categorical differences in RTs being due to differences in valence ambivalence (Cunningham, Raye, & Johnson, 2004) - as participants might take longer at considering both positive and negative features - we also administered an additional online survey to 40 participants (33 females, age range: 21-35). The survey presented participants with the target stimuli in a random order and asked to rate on a 9-point Likert scale the degree of both the positivity and negativity of each concept (0 = not at all, 9 = very much).To calculate ambivalence from the survey data, we computed for each concept an index of ambivalence using the equation of the Gradual Threshold Model (Priester & Petty, 1996) and subjected it to a 2 (Category: Social vs. Nonsocial) x 2 (Valence: Positive vs. Negative) ANOVA. The analysis of the resulting ambivalence scores revealed no effect of Category [$F(1, 36) = .60$, $p = .44$, $\eta_p^2 = .02$], a main effect of Valence [$F(1, 36) = 44.71$, $p < .001$, $\eta_p^2 = .55$], with negative concepts being overall more ambivalent, and no significant interaction [$F(1, 36) = 3.03$, $p = .09$, $\eta_p^2 = .08$].

A sample of 40 pictures representing the same concepts expressed by the selected words was retrieved from Google Images and matched across category and valence polarity in their representativeness ratings, collected from a different sample of 14 participants (see Table 1). Where applicable, the faces of the depicted individuals were manually blurred using the GIMP 2 software (Kimball, Mattis, Natterer, & Neumann, 2013). Additionally, pictures were also matched in low-level visual features (all *p*s > .05). These included average luminance, amount of red, green and blue, and average spatial frequency, extracted through a custom Matlab script (adapted from Blechert, Meule, Busch, & Ohla, 2014).

Variables	Valence	Arousal	Familiarity	Length	Ambivalence	Pic. represent.
Social Positive	6.14 ± .71	3.91 ± .62	3.46 ± 1.02	7.4 ± 1.17	1.82 ± 2.06	6.69 ± .26
Object Positive	6.42 ± .47	4.00 ± .69	3.8 ± 1.04	6.8 ± 1.48	0.45 ± 1.01	6.64 ± .16
Social Negative	3.30 ± .81	3.88 ± .57	3.63 ± .75	7.9 ± .52	4.50 ± 2.10	6.64 ± .27
Object Negative	3.07 ± .48	3.81 ± .57	2.94 ± 1.13	7.1 ± 1.52	5.03 ± 1.44	6.70 ± .20
ANOVA	$p < .001$	$p = .92$	$p = .26$	$p = .37$	$p = .09$	$p = .87$

Table 1. Descriptive statistics for selected words and pictures, divided the combination of semantic category and valence polarity. Means of word length and ratings of valence (1, unpleasant to 7, pleasant), arousal (1, not arousing to 7, very arousing), familiarity (1, unfamiliar to 7, familiar), ambivalence, and of picture representativeness (1, not representative to 7, very representative). On the bottom row, *p*-values of one-way ANOVAs on each factor, using as independent variable the combination of category and valence polarity.

To ensure an equal speed in semantic access between categories, we collected a measure of accessibility by devising a simple categorisation task. In this task, an independent sample of 20 participants (11 females, age range: 25-35) was asked to provide speeded responses to the selected stimuli. The task required to categorize the targets based on whether they represented people or objects, using the same timings of the main task (see Fig. 1). The analysis of the reaction times showed no main effect of category [$F(1, 19) = .53$, $p = .48$, $\eta_p^2 = .03$], for which participants took on average the same time to categorize social groups and objects, regardless of modality and valence (both *p*s > .05).

2.3 Procedure

Following the montage of the EEG cap, participants were seated in an acoustically isolated room and asked to fixate an “X” in the center of the screen for 3 minutes to record resting-state EEG activity.

After this period, they completed an evaluative requiring to indicate via button press whether the concept represented by the target stimulus was associated with a pleasant or unpleasant feeling. A PC running PsychoPy (Peirce, 2007) controlled the presentation of the stimuli and the recording of responses. Stimuli were projected on a gray background via a 19” LCD monitor with resolution of 1280*1024 pixels and a screen refresh rate of 60 Hz. During each trial, a fixation cross was presented for 700 ms, followed by a 200 ms blank screen and by the presentation of the target stimulus (300 ms). The target was then replaced by another blank screen that lasted for 1700 ms, giving participants a total of 2000 ms to respond from target onset (see Figure 1). The evaluation was given via button press by using the index finger of each hand placed over the ‘f’ and ‘j’ QUERTY keyboard buttons. The intertrial interval (ITI) was jittered between 800 and 1200 ms at 100 ms intervals, and presented a white fixation cross on a gray background. Participants were instructed to try to restrict their blinking to the ITI period to reduce the number of artefacts. The target stimuli consisted of images or words representing social groups or objects of positive or negative valence, presented in a random order. Participants first completed a block of 10 practice trials, followed by 12 test blocks of 40 trials each. Single blocks contained either pictures of words in an alternating, counterbalanced order, and were separated by self-paced breaks.

The experiment consisted of a 2 (Category: social group vs. object) x 2 (Valence: positive vs. negative) x 2 (Modality: picture vs. word) within-subjects design, with response time (RT), accuracy, and EEG voltage as the dependent variables.

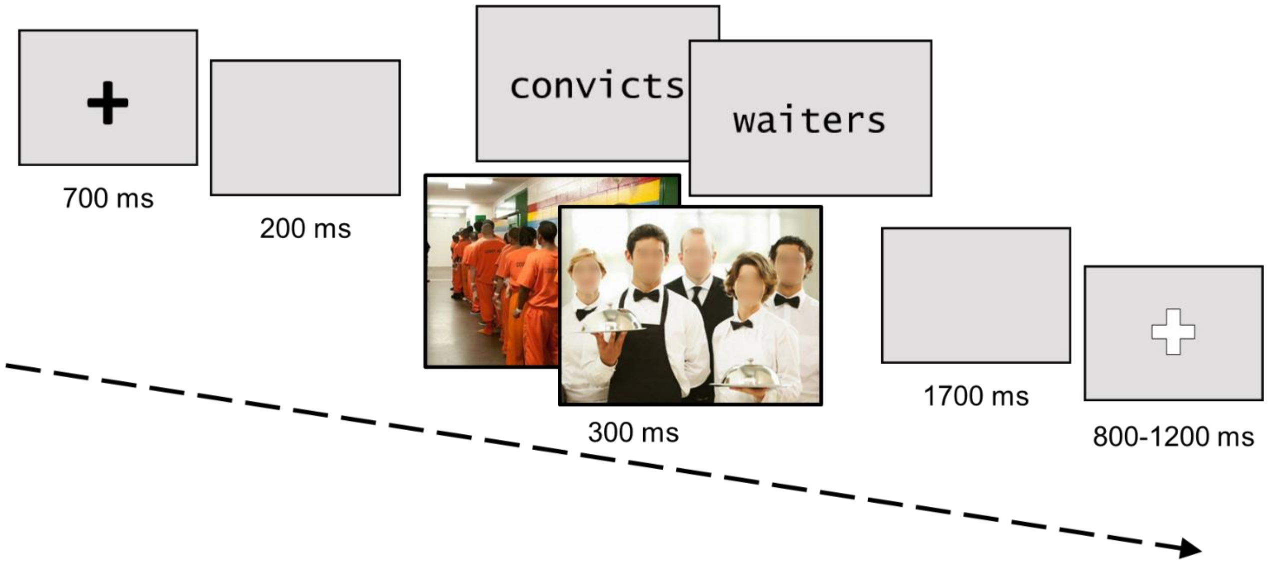

Figure 1. Temporal progression an evaluation trial with examples of social groups of negative and positive valence, represented as words or pictures (only one of the stimuli was presented each time). Participants were instructed to evaluate the target starting from its onset, and to try not to blink until the appearance of the white cross (ITI).

2.4 Electrophysiological recordings

A set of 64 Ag/AgCl active electrodes connected to a BioSemi Active-Two amplifier system were mounted on an elastic cap according to the International 10/20 system to record the continuous neural signal by means of ActiView acquisition software (Biosemi, Amsterdam, Netherlands). Electrode offsets were kept between ±20 mV, while the signal was sampled at a rate of 1024 Hz with a 24-bit resolution. A common mode voltage based on the ActiveTwo's CMS/DRL feedback loop was used for analog-to-digital conversion of recorded voltages for each electrode (cf. to https://www.biosemi.com/faq/cms%26drl.htm). Anti-aliasing filters were used and data were band-pass filtered between 0.01–100 Hz during data acquisition.

EEG data preprocessing was performed using the Brainstorm software (Tadel, Baillet, Mosher, Pantazis, & Leahy, 2011). First, the EEG recordings were downsampled offline at 125 Hz and band-pass filtered (0.05 – 40 Hz). Bad electrode channels were removed upon visual inspection, as well as movement artefacts. Eye blinks artefacts were removed through independent components analysis (ICA; (Makeig, Bell, Jung, & Sejnowski, 1996). Epochs containing artefacts other than eye movements were removed after visual inspection. One subject (female, age 24) was excluded from all further analyses due to the presence of excessive movement artefacts. The data were then epoched from -200 to +1500 ms relative to the onset of the target and baseline-corrected from -200 ms to target onset.

2.5 Behavioral data analysis

Behavioral data were preprocessed and analyzed in R (R Core Team, 2016). Data from correct trials within 2 SD from each participant's average were then log-transformaed to reduce the skew of the distribution and subjected to a 2 (Category: Social vs. Nonsocial) x 2 (Valence: Positive vs. Negative) x 2 (Modality: names vs. pictures) repeated-measures ANOVA. Follow-ups to significant interactions consisted of Bonferroni-corrected contrasts. Only results involving interactions with the Category factor and yielding significant post-hoc contrasts are reported.

2.6 ERP data analysis

The time windows for the main components were identified from the existing literature (Hinojosa et al., 2009; Zhang et al., 2014; Zhu et al., 2015) and the visual inspection of average peak amplitudes. In order not to lose statistical power in quantifying effects over several electrodes, eleven regions of interest (ROIs; see Figure 2) were then created by averaging the amplitude of the respective electrodes (as in Hinojosa, Carretié, Valcárcel, Méndez-Bértolo, & Pozo, 2009). The resulting ERPs were subjected to two 2 (Category: Social vs. Nonsocial) x 2 (Valence: positive vs. negative) x 2 (Modality: word vs. picture) x 11 (Cluster: [OC, LP, RP, LC, RC, LF, RF, FP, MF, MC, MP]) repeated measures ANOVAs, one for each component's time window (150 - 200 ms, 400 - 700 ms and 700 - 1000 ms). Where necessary, the degrees of freedom of the *F* ratios were adjusted using the Greenhouse–Geisser epsilon correction. Follow-ups to significant interactions including Valence consisted of Bonferroni-corrected contrasts of Valence effects. Only interactions yielding significant follow-up contrasts are reported.

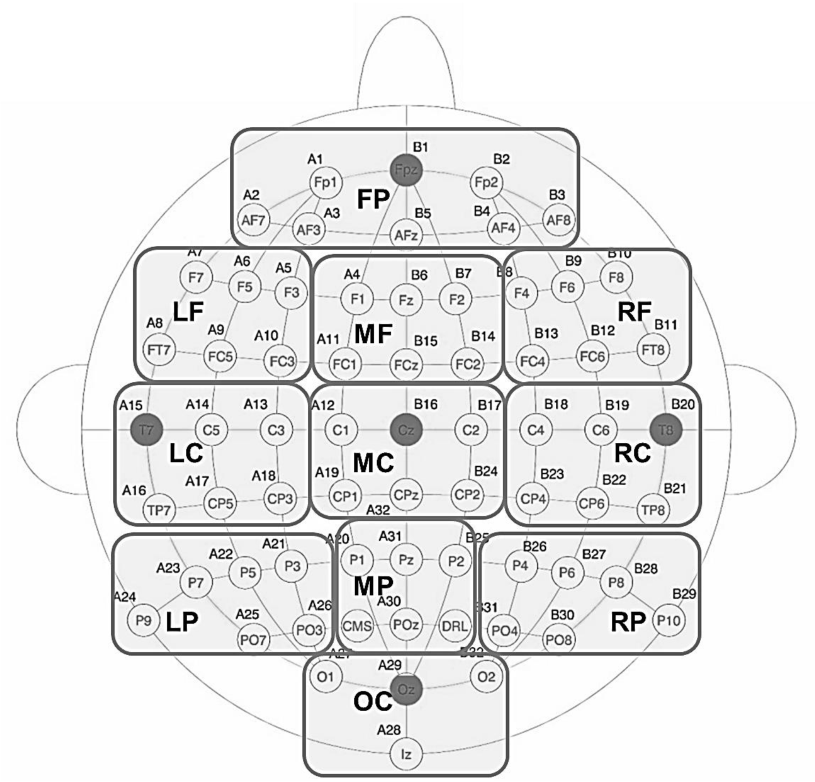

Figure 2. Clusters of electrodes grouped for statistical analysis of ERPs. OC, occipital, LP, left posterior, RP, right posterior, LC, left central, RC, right central, LF, left frontal, RF, right frontal, FP, frontopolar, MF, middle frontal, MC, middle central, MP, middle posterior.

2.7 Decoding analysis

Multivariate classification analyses were performed using the CoSMoMVPA analysis package (www.cosmomvpa.org) (Oosterhof, Connolly, & Haxby, 2016) implemented in MATLAB. Classification was performed separately for every 8 ms time bin using linear discriminant analysis (LDA) classifiers. These were trained to discriminate the patterns of activation across EEG sensors for the two valence conditions in one subset of the data, and tested on another.

Cross-decoding was conducted separately for Social and Object categories. Training data consisted of trials from the Words condition, after which classifiers were tested on trials from the Pictures condition. In this way, the decoding performed by the classifier on pictures derived from the identification of the same patterns of valence differences that had been learned on word stimuli. The correctly identified patterns were thus common to both modalities. To increase the reliability signal-to-noise ratio and the of the data for the classifier, for each participant and experimental condition separately, five averaged trials were created. As the number of correct trials differed between participants and within conditions, the averages containing roughly the trials corresponding to each experimental block/run, with the constraint that no average was derived from more than one trial more than the other averages.

Given the consistency of the response mappings within participants, to avoid the confound of decoding motor responses rather than valence differences (Grootswagers et al., 2017), classifiers were trained and tested on the combined data from pairs of participants with opposite mappings. The classification accuracy of each participant was calculated by averaging the performance of all the classifications containing that participant's data (i.e., 10 for each subject). The percentage of correct predictions of the classifier was used as index of classification accuracy. The classification was generalized across train and test times, for which it was repeated for their every possible combination, leading to a classification accuracy map of 125 x 125 points (i.e., 1000 ms x 1000 ms with 125 Hz resolution) for every comparison in each participant. Individual maps were smoothed with an averaging box filter of the size of 3 x 3 time points (i.e., 24 ms in both training and testing time).

2.8 Statistical testing

To identify time-periods presenting above chance classification accuracy, we used threshold-free cluster-estimation procedure (Smith & Nichols, 2009) with default parameters from the CoSMoMVPA package (Oosterhof et al., 2016), using multiple comparison correction based on a sign-permutation test (with null distributions created from 10,000 bootstrapping iterations). To reveal where decoding performance was significant, the threshold on the statistical maps was set at $Z > 2.57$ (i.e., $p < .005$). The procedure was first applied within each category against the mean of 0.5 decoding accuracy and then used to check for significant differences in contrasting the two.

3. Results

3.1 Behavioral results

Reaction times analysis showed a main effect of category [$F(1, 19) = 128.83$, $p < .001$, $\eta_p^2 = .87$] with faster responses to non-social than to social categories. No other significant main effects were present. Significant first order interactions emerged between category and valence [$F(1, 19) = 18.24$, $p < .001$, $\eta_p^2 = .49$], and between category and modality [$F(1, 19) = 47.94$, $p < .001$, $\eta_p^2 = .72$], with the former indicating how the difference in RTs between social groups and objects was significantly greater in the positive relative to the negative valence targets ($p = .001$), and in pictures relative to words ($p < .001$). Last, a post-hoc analysis of the significant second-order interaction between category, valence and modality [$F(1, 19) = 8.37$, $p = .009$, $\eta_p^2 = .31$], revealed that social groups were evaluated slower relative to objects in all valence and modality combinations (all ps $< .001$) but the negative word one, where there was no significant categorical difference ($p = .60$). (see Figure 3).

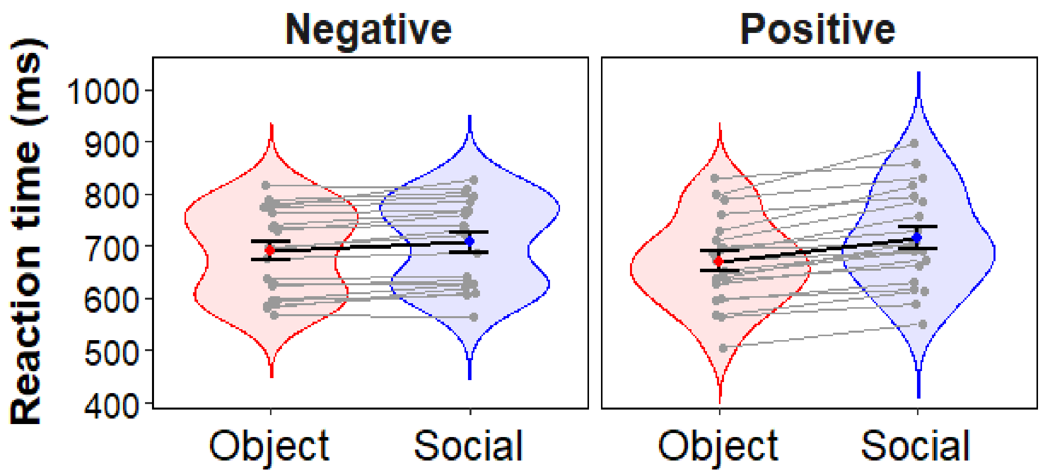

Figure 3. Violin plots displaying single participant mean RT distribution and group mean RTs for Object and Social category targets divided by valence. Error bars represent ± 1 SE.

3.2 ERP results

150 – 200 ms. The analysis of the P1 time window resulted in a significant three-way interaction between Valence, Category and Cluster [$F(3.19, 60.53) = 3.28$, $p = .025$, $\eta_p^2 = .15$], and a significant four-way interaction between Valence, Category, Modality and Cluster [$F(2.61, 49.61) = 7.54$, $p < .001$, $\eta_p^2 = .28$]. Follow-up analyses for the word modality evidenced a significant Valence effect in the right frontal cluster ($p < .05$), with a greater amplitude for negative as opposed to positive stimuli ($p < .05$). The effect was only significant for Social groups ($p < .05$), while the effect for and difference from Objects was not significant (both *p*s > .1). No effects were present for pictures in the same cluster. For pictures, a significant Valence effect was present for Objects in the occipital, left posterior and left anterior clusters (*p*s < .05), with no effect for Social groups (*p*s > .05), and a significant difference between categories (*p*s < .05). A Valence effect for both categories was present in the right posterior and the mid frontal clusters (*p*s < .05), presenting also a significant difference between categories ($p < .05$). This difference was characterized by a greater amplitude for negative vs positive stimuli for Objects and a greater amplitude for positive vs negative stimuli for Social groups.

400 – 700 ms. A significant interaction was present between Valence, Category and Cluster [$F(3.71, 70.49) = 3.27$, $p = .02$, $\eta_p^2 = .15$]. At the single cluster level, a marginally significant Valence effect, with a greater amplitude for positive stimuli, was only present for words in the central right electrodes ($p = .07$). In the picture modality, the same cluster also presented a significant Valence effect for Objects ($p < .05$), where the difference with Social groups was marginally significant ($p = .08$).

700 – 1000 ms. Significant interactions between Valence, Modality and Cluster [$F(3.84, 72.98) = 2.54$, $p = .049$, $\eta_p^2 = .12$], Valence, Category and Cluster [$F(4.36, 82.75) = 2.62$, $p = .036$, $\eta_p^2 = .12$], Valence, Category and Modality [$F(1, 19) = 6.87$, $p = .017$, $\eta_p^2 = .27$] were present in the latest time window. These were accompanied by a marginally significant four-way interaction involving

Valence, Category, Modality and Cluster [$F(3.24, 61.57) = 2.51$, $p = .063$, $\eta_p^2 = .12$]. Follow-up analyses showed a significant Valence effect in the right posterior cluster for words ($p < .05$) but not for pictures ($p > .1$). Within this cluster, the effect of Valence was significant for Social groups ($p < .05$) but not for Objects ($p > .1$), with the difference between the two categories also being significant ($p < .05$). A marginally significant Valence effect in the left central cluster for words only ($p = .07$), with a greater amplitude for negative stimuli, was driven by a significant effect for Social groups ($p < .05$), with no effect for Objects and no significant difference between categories (both *p*s > .1). In the right central cluster, the main effect of Valence ($p < .05$) was accompanied by two marginally significant effects in both categories (*p*s = .07), in all cases displaying a greater amplitude for positive relative to negative word stimuli. For pictures, only a marginally significant Valence effect emerged for Social groups in the left posterior region ($p = .09$).

3.1 Cross-modal valence decoding in social groups and object categories

Valence cross-decoding results for social groups evidenced a first significant decoder performance period at 100-150 ms, followed by a second interval going from ~450 to 1000 ms (see Figure 3A & 3C). In the case of object categories, only one significant window was present between ~400 and 700 ms (see Figure 3B & 3C). When contrasting decoder performance between categories, a significantly greater decoding performance was found for social groups in two time windows - an early window going from 50 to 150 ms and a late window starting at 750 ms up to 1000 ms (see Figure 3D).

Time-generalisation results evidenced how decoding did not differ based on modality, as all significant regions laid on the diagonal (see Figures 4A & 4B). Additionally, a sustained pattern of neural activity across modalities was present for social groups in the ~450-1000 ms time window (King & Dehaene, 2014), as evidenced by the square-shaped region of significance (see Figure 4A). This pattern was unique to social groups and was not present for objects (see Figure 4B).

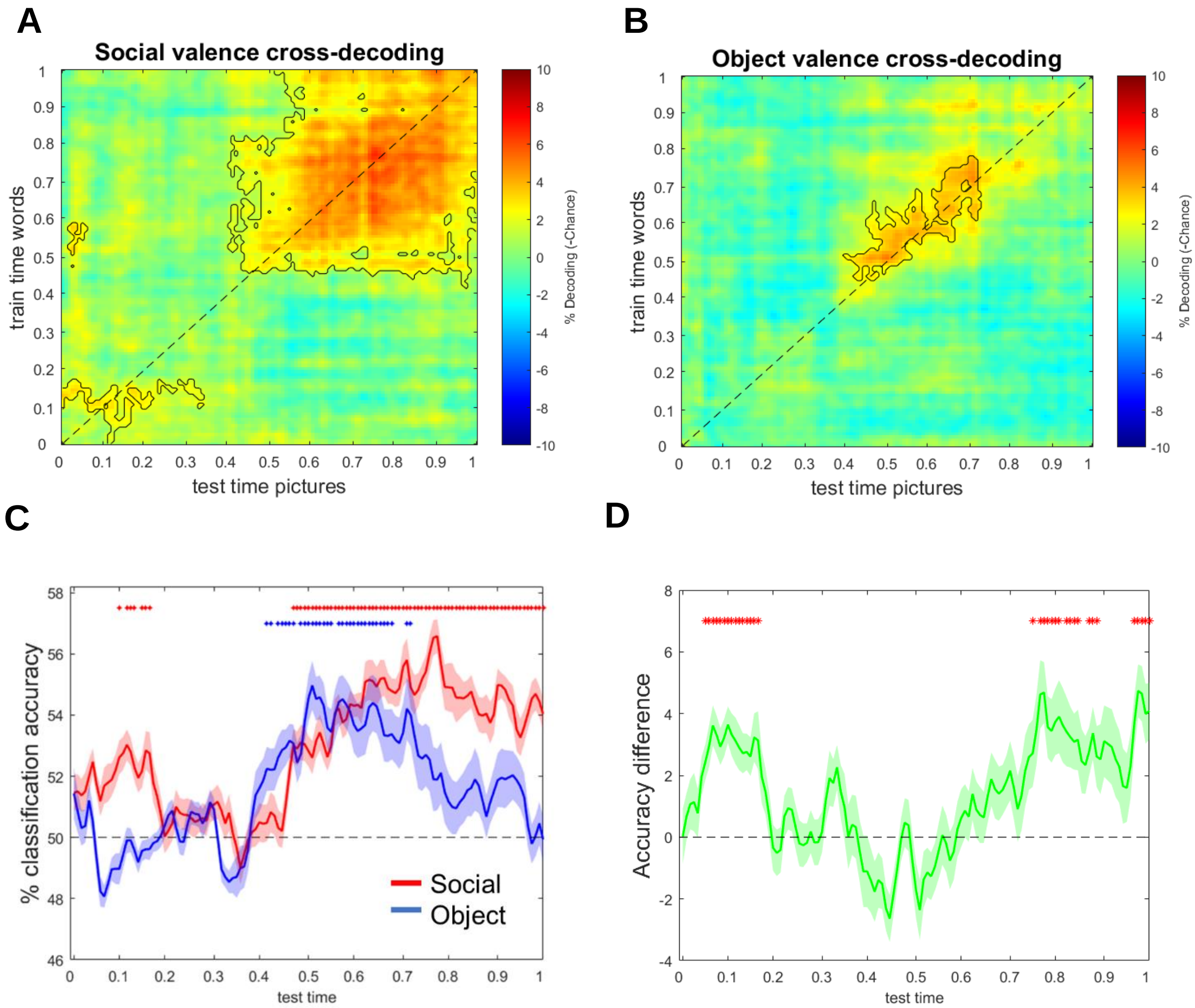

Figure 4. A) Time-generalization plots resulting from cross-decoding valence of social category and object category targets **(B)**. Black contours represent above chance decoding accuracy (*p* < .001). **C)** Classification accuracy across the time-generalization diagonal for social and object category targets, and **(D)** their difference. Asterisks indicate above chance decoding accuracy time intervals (*p* < .001).

4. **Discussion**

With the present study we aimed at identifying the distinctive patterns of brain activity associated with the affective processing of social groups and compare it to the one of nonsocial categories. Following previous neuropsychological studies (Piretti et al., 2015; Rumiati et al., 2014; Suran, Rumiati, et al., 2019) and theoretical propositions suggesting a greater relevance of affective features in representing social categories (Lambon Ralph et al., 2016; Mahon & Caramazza, 2011; Simmons & Barsalou, 2003), and following studies showing both common and distinct anatomical bases for processing social and non-social affective information (Norris et al., 2004), we expected to find categorical differences in the temporal dynamics of affective processing, prioritizing the decoding of valence in social groups compared to nonsocial categories. To rule out possible confounds generated by the input modality and thus isolate affective evaluation, we cross-decoded the effects between two different input-modalities (words and images) using MVPA by training of classifiers using neural data associated with one modality (e.g. words) and test on data from the other modality (pictures). Additionally, we applied time generalization to study the evolution of the affect-specific neural code while locating its processing in the temporal dimension of the two input modalities. Our results showed significant differences in decoding affective features between social and nonsocial category targets. This difference was present in both early and later time windows, and was also reflected in overall response times as longer responses to social groups, but just when participants were required to make affective judgments.

Significant decoding performance appeared first at 100 ms post-stimulus that was unique to social categories. This earlier decoding might indicate the prioritization of affective processing for Social groups over Objects. This effect might also be linked to the early effects found by previous ERP studies using words (Zhang et al., 2014) and pictures (Zhu et al., 2015) which, although here not detected across modalities by univariate analyses, was captured by the greater sensitivity and earlier detection attributed to MVPA (Grootswagers et al., 2017). In the second time window of significant decoding, above-chance accuracy started approximately at ~400 ms for nonsocial categories and at ~450 ms for social groups, reflecting the delay found in the behavioral response. This time interval overlaps with the emergence of the late positive components modulated by the affective content of both words and pictures (Zhang et al., 2014; Zhu et al., 2015), and it represents the only time interval in which a significant difference between negative and positive stimuli was present across modalities in the same scalp region when analysing ERPs. While for Objects significant decoding in the second time window lasted until ~700 ms, significant affective decoding in Social groups extended to 1000 ms post-stimulus and was significantly greater than the one for Object categories from ~750 ms on. In previous ERP research, effects occurring in this time window are associated with a greater sustained processing of emotional stimuli (Citron, 2012) and an enhanced motivational significance (Liu, Huang, McGinnis-Deweese, Keil, & Ding, 2012; Schupp et al., 2000; Tempel et al., 2013), which might be intrinsically greater in social groups due to the presence of conspecifics. Although more valence effects were identified with the univariate analysis for Social groups relative to Objects, none of these were captured across modalities and in the same regions. The claim of a more sustained processing of Social groups is also supported by the results of the time generalization, presenting a square-shaped pattern of significance, indeed associated with sustained patterns of brain activity in the decoded dimension (King & Dehaene, 2014).

Reaction times to social groups evidenced a slower evaluative response in comparison to nonsocial categories. The lack of such differences in ambivalence and categorization times of the same targets allows us to exclude the possibility that the disparity was due to social group stimuli presenting stronger conflicting (i.e., nondominant) affective associations (Cunningham, Johnson, Gatenby, Gore, & Banaji, 2003; Cunningham et al., 2004), or having a delayed semantic access, respectively. As suggested by previous neuroimaging studies evidencing slower responses in categorizing phrases describing social interactions versus nonsocial actions (Wood, Romero, Makale, & Grafman, 2003), we speculate that it is possible that participants activated more complex representations for social relative to nonsocial categories to make affective decisions. Evaluative processes generally involve a more complex selection of information in comparison to categorization (Cunningham et al., 2003), thus task demands might interact with the sociality of the semantic categories of target stimuli. For example, the present data might reflect how the evaluation of social groups requires accessing all relevant representations of the behaviors commonly associated with them, while in the case of nonsocial categories, just the representation of their common usage may have sufficed.

In conclusion, our results suggest that processing of affective features comes with distinct temporal dynamics for social groups and nonsocial categories, and represent further evidence of the interaction between semantic and affective information in the temporal domain. An early decoding of affective features for social groups however was not associated with faster responses, as the behavioural findings have shown that subjects took longer in providing a behavioural response when evaluating social groups as opposed to objects. Interestingly, this effect seems to be independent from the timing of access to the more basic person/object categorical representations. It is indeed possible that a slower evaluative response to social groups was triggered by self-presentation concerns when required to provide judgments about other people following access to the categorical semantic information (Nosek, 2005), and thus requiring additional affective processing stages as opposed to objects. The same sustained processing of affective information of social groups was also evidenced at a neural level, with longer lasting classifier decoding accuracy in a late time window, going beyond the behavioural response. We suggest that the delayed response and longer decoding of affective features in social groups, possibly associated with more sustained post-semantic attentional processes due to more complex representations, supports the hypothesis of a greater relevance of their affective features compared to nonsocial categories (Rumiati et al., 2014). As our conclusions are derived from a setting explicitly requiring subjects to focus only on affective content, it is possible that its increased weight was specific to the present task demands. By directly comparing the neural signatures resulting from different tasks, future studies may focus on elucidating which aspects of affective feature processing in social groups interact with experimental demands and which ones occur independently from it, being thus more intrinsic to their processing.

REFERENCES

Blechert, J., Meule, A., Busch, N. A., & Ohla, K. (2014). Food-pics: An image database for experimental research on eating and appetite. *Frontiers in Psychology*, *5*(JUN). https://doi.org/10.3389/fpsyg.2014.00617

Citron, F. M. M. (2012). Neural correlates of written emotion word processing: A review of recent electrophysiological and hemodynamic neuroimaging studies. *Brain and Language*. https://doi.org/10.1016/j.bandl.2011.12.007

Cunningham, W. A., Espinet, S. D., Deyoung, C. G., & Zelazo, P. D. (2005). Attitudes to the right- and left: Frontal ERP asymmetries associated with stimulus valence and processing goals. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2005.04.044

Cunningham, W. A., Johnson, M. K., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2003). Neural components of social evaluation. *Journal of Personality and Social Psychology*, *85*(4), 639–649. https://doi.org/10.1037/0022-3514.85.4.639

Cunningham, W. A., Raye, C. L., & Johnson, M. K. (2004). Implicit and Explicit Evaluation: fMRI Correlates of Valence, Emotional Intensity, and Control in the Processing of Attitudes. *Journal of Cognitive Neuroscience*, *16*(10), 1717–1729. https://doi.org/10.1162/0898929042947919

Fairhall, S. L., & Caramazza, A. (2013). Brain Regions That Represent Amodal Conceptual Knowledge. *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.0051-13.2013

Grootswagers, T., Wardle, S. G., & Carlson, T. A. (2017). Decoding dynamic brain patterns from evoked responses: A tutorial on multivariate pattern analysis applied to time series neuroimaging data. *Journal of Cognitive Neuroscience*. https://doi.org/10.1162/jocn_a_01068

Hinojosa, J. A., Carretié, L., Valcárcel, M. A., Méndez-Bértolo, C., & Pozo, M. A. (2009).

Electrophysiological differences in the processing of affective information in words and pictures. *Cognitive, Affective, & Behavioral Neuroscience*, *9*(2), 173–189. https://doi.org/10.3758/CABN.9.2.173

Kimball, S., Mattis, P., Natterer, M., & Neumann, S. (2013). *Gimp: Gnu image manipulation program.* Retrieved from www.gimp.org

King, J. R., & Dehaene, S. (2014). Characterizing the dynamics of mental representations: The temporal generalization method. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2014.01.002

Lambon Ralph, M. A., Jefferies, E., Patterson, K., & Rogers, T. T. (2016). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience*, *18*(1), 42–55. https://doi.org/10.1038/nrn.2016.150

Leonardelli, E., Fait, E., & Fairhall, S. L. (2019). Temporal dynamics of access to amodal representations of category-level conceptual information. *Scientific Reports*, *9*(1), 239. https://doi.org/10.1038/s41598-018-37429-2

Liu, Y., Huang, H., McGinnis-Deweese, M., Keil, A., & Ding, M. (2012). Neural Substrate of the Late Positive Potential in Emotional Processing. *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.3109-12.2012

Mahon, B. Z., & Caramazza, A. (2011). What drives the organization of object knowledge in the brain? *Trends in Cognitive Sciences*, *15*(3), 97–103. https://doi.org/10.1016/j.tics.2011.01.004

Makeig, S., Bell, A. J., Jung, T.-P., & Sejnowski, T. J. (1996). Independent Component Analysis of Electroencephalographic Data. *Advances in Neural Information Processing Systems*, *8*, 145–151. https://doi.org/10.1109/ICOSP.2002.1180091

Norris, C. J., Chen, E. E., Zhu, D. C., Small, S. L., & Cacioppo, J. T. (2004). The Interaction of

Social and Emotional Processes in the Brain. *Journal of Cognitive Neuroscience*, *16*(10), 1818–1829. https://doi.org/10.1162/0898929042947847

Nosek, B. A. (2005). Moderators of the relationship between implicit and explicit evaluation. *Journal of Experimental Psychology: General*. https://doi.org/10.1037/0096-3445.134.4.565

Oosterhof, N. N., Connolly, A. C., & Haxby, J. V. (2016). CoSMoMVPA: Multi-Modal Multivariate Pattern Analysis of Neuroimaging Data in Matlab/GNU Octave. *Frontiers in Neuroinformatics*. https://doi.org/10.3389/fninf.2016.00027

Peirce, J. W. (2007). PsychoPy-Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1–2), 8–13. https://doi.org/10.1016/j.jneumeth.2006.11.017

Piretti, L., Carnaghi, A., Campanella, F., Ambron, E., Skrap, M., & Rumiati, R. I. (2015). The neural network associated with lexical-semantic knowledge about social groups. *Cortex*, *70*, 155–168. https://doi.org/10.1016/j.cortex.2015.06.024

Priester, J. R., & Petty, R. E. (1996). The Gradual Threshold Model of Ambivalence: Relating the Positive and Negative Bases of Attitudes to Subjective Ambivalence. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/0022-3514.71.3.431

R Core Team. (2016). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria.

Rumiati, R. I., Carnaghi, A., Improta, E., Diez, A. L., & Silveri, M. C. (2014). Social groups have a representation of their own: Clues from neuropsychology. *Cognitive Neuroscience*, *5*(2), 85–96. https://doi.org/10.1080/17588928.2013.876981

Schupp, H. T., Cuthbert, B. N., Bradley, M. M., Cacioppo, J. T., Tiffany, I., & Lang, P. J. (2000). Affective picture processing: The late positive potential is modulated by motivational relevance. *Psychophysiology*. https://doi.org/10.1017/S0048577200001530

Simmons, W. K., & Barsalou, L. W. (2003). The similarity-in-topography principle: reconciling theories of conceptual deficits. *Cognitive Neuropsychology*, *20*(3–6), 451–486. https://doi.org/10.1080/02643290342000032

Smith, S. M., & Nichols, T. E. (2009). Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2008.03.061

Suran, T., Arcara, G., Piretti, L., & Rumiati, R. I. (2019). *Electrophysiological Correlates of Social Group Representations in Affective Priming*. https://doi.org/10.31234/osf.io/kewu5

Suran, T., Rumiati, R. I., & Piretti, L. (2019). The contribution of the left inferior frontal gyrus in affective processing of social groups. *Cognitive Neuroscience*, 1–10. https://doi.org/10.1080/17588928.2019.1593127

Tadel, F., Baillet, S., Mosher, J. C., Pantazis, D., & Leahy, R. M. (2011). Brainstorm: A User-Friendly Application for MEG/EEG Analysis. *Computational Intelligence and Neuroscience, 2011*, 1–13. https://doi.org/10.1155/2011/879716

Tempel, K., Kuchinke, L., Urton, K., Schlochtermeier, L. H., Kappelhoff, H., & Jacobs, A. M. (2013). Effects of positive pictograms and words: Anemotional word superiority effect? *Journal of Neurolinguistics*. https://doi.org/10.1016/j.jneuroling.2013.05.002

Wood, J. N., Romero, S. G., Makale, M., & Grafman, J. (2003). Category-Specific Representations of Social and Nonsocial Knowledge in the Human Prefrontal Cortex. *Journal of Cognitive Neuroscience*, *15*(2), 236–248. https://doi.org/10.1162/089892903321208178

Zhang, D., He, W., Wang, T., Luo, W., Zhu, X., Gu, R., … Luo, Y. (2014). Three stages of emotional word processing: an ERP study with rapid serial visual presentation. *Social Cognitive and Affective Neuroscience*, *9*(12), 1897–1903. https://doi.org/10.1093/scan/nst188

Zhu, C., He, W., Qi, Z., Wang, L., Song, D., Zhan, L., … Luo, W. (2015). The time course of emotional picture processing: an event-related potential study using a rapid serial visual presentation paradigm. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2015.00954

APPENDIX

Target word stimuli [English translation]

Social groups		Objects	
Negative	**Positive**	**Negative**	**Positive**
alcolisti [alcoholics]	bambini [children]	bare [coffins]	anelli [rings]
barboni [homeless]	anziani [elderly]	coltelli [knives]	biciclette [bicycles]
detenuti [convicts]	suore [nuns]	droghe [drugs]	chitarre [guitars]
drogati [junkies]	camerieri [waiters]	fucili [rifles]	cuscini [pillows]
migrant [migrants]	marina [sailors]	granate [grenades]	libri [books]
militari [military]	ballerina [dancers]	manette [handcuffs]	matite [pencils]
obesi [obese]	bagnini [lifeguards]	pistole [guns]	orology [watches]
prostitute [prostitutes]	pompieri [firefighters]	sigarette [cigarettes]	palloni [balloons]
terroristi [terrorists]	scultori [sculptors]	siringhe [syringes]	perle [pearls]
zingari [gypsies]	pittori [painters]	stampelle [crutches]	violini [violins]